RFLP analysis

RFLP= Restriction fragment length polymorphism

- Refers to variation in restriction sites between individuals in a population
- These are extremely useful and valuable for geneticists (and lawyers)
- On average two individuals (humans) vary at 1 in 1000 bp
- The human genome is $3x10^9$ bp
- This means that they will differ in more than 3 million bp.
- By chance these changes will create or destroy the recognition sites for Restriction enzymes

RFLP

Lets generate a restriction map for a region of human X-chromosome

The restriction map in the same region of the X chromosome of a second individual may appear as

Normal ***GAATTC***

Mutant ***GAGTTC***

RFLP

The internal EcoRI site is missing in the second individual

For X1 the sequence at this site is `GAATTC`
`CTTAAG`

This is the sequence recognized by EcoRI

The equivalent site in the X2 individual is mutated

`GAGTTC`
`CTCAAG`

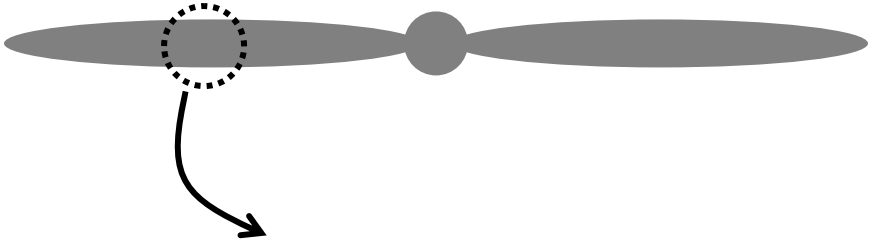
Now if we examine a large number of humans at this site we may find that 25% possess the EcoRI site and 75% lack this site.

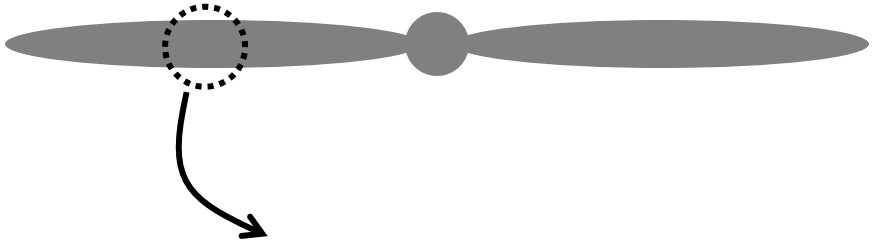
We can say that a restriction fragment length polymorphism exits in this region

These polymorphisms usually do not have any phenotypic consequences
Silent mutations that do not alter the protein sequence because of redundancy in Codon usage, localization to introns or non-genic regions or do not affect protein Structure/function.

RFLP

RFLP are identified by southern blots
In the region of the human X chromosome, two forms of the X-chromosome are Segregating in the population.





Digest DNA with EcoRI and probe with probe1
What do we get?

RFLP

Digesting with BamHI and performing Southern blots with the above probe produces the following results for X1/X1, X1/X2 and X2/X2 individuals:

There is no variation with respect to the BamHI sites, all individuals produce the same banding patterns on Southern blots

If we used probe2 for southern blots with a BamHI digest what would be the Results for X1/X1, X1/X2 and X2/X2 individuals?

If we used probe2 for southern blots with a EcoRI digest what would be the results for X1/X1, X1/X2 and X2/X2 individuals?

RFLP

RFLP's are found by trial and error and they require an appropriate probe and enzyme They are very valuable because they can be used just like any other genetic marker to map genes

They are employed in recombination analysis (mapping) in the same way as conventional Allelic variants are employed

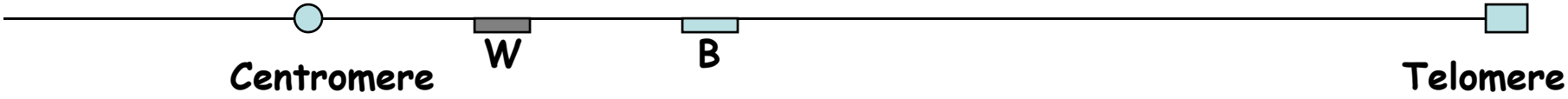
The presence of a specific restriction site at a specific locus on one chromosome and its absence at a specific locus on another chromosome can be viewed as two allelic forms of a gene

The phenotype in this case is a Southern blot rather than white eye/red eye

Lets review standard mapping:
To map any two genes with respect to one another, they must be heterozygous at both loci.

Mapping

Gene W and B are responsible for wing and bristle development



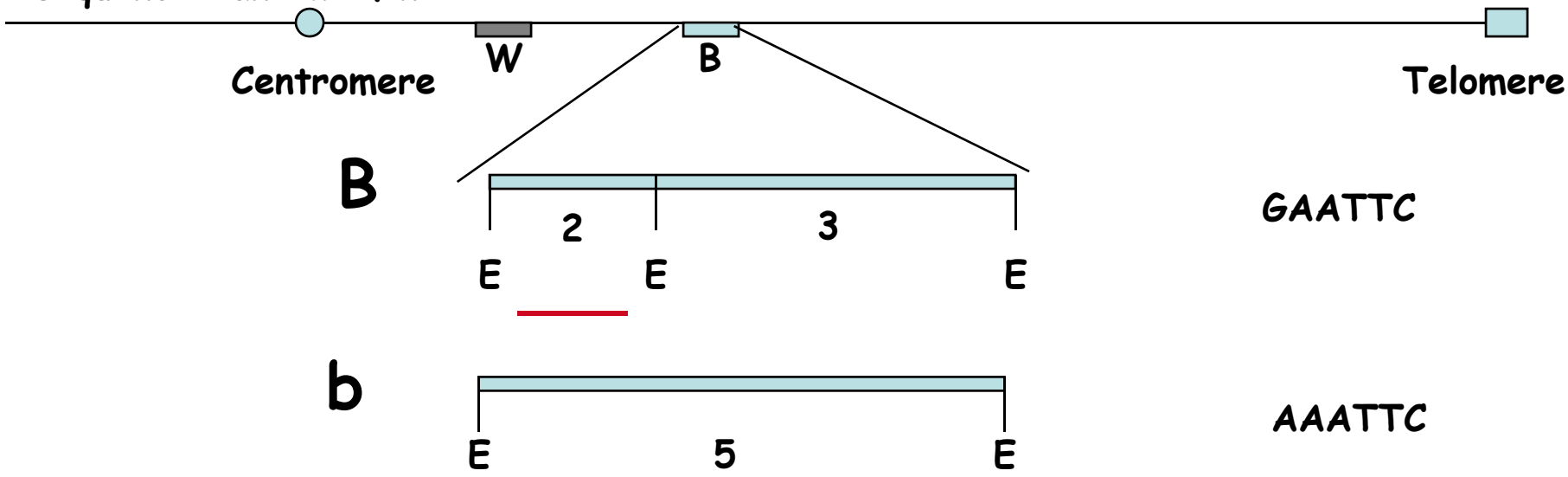
To find the map distance between these two genes we need allelic variants at each locus

W=wings	B=Bristles
w= No wings	b= no bristles

To measure genetic distance between these two genes, the double heterozygote is crossed to the double homozygote

Mapping

Both the normal and mutant alleles of gene B (B and b) are sequenced and we find



By chance, this mutation disrupts the amino acid sequence and also a EcoRI site!

If DNA is isolated from B/B, B/b and b/b individuals, cut with EcoRI and probed in A Southern blot, the pattern that we will obtain will be

B/B Bristle B/b Bristle b/b No bristle

Mapping

Therefore in the previous cross (WB/wb x wb/wb), the genotype at the B locus can be distinguished either by the presence and absence of bristles or Southern blots

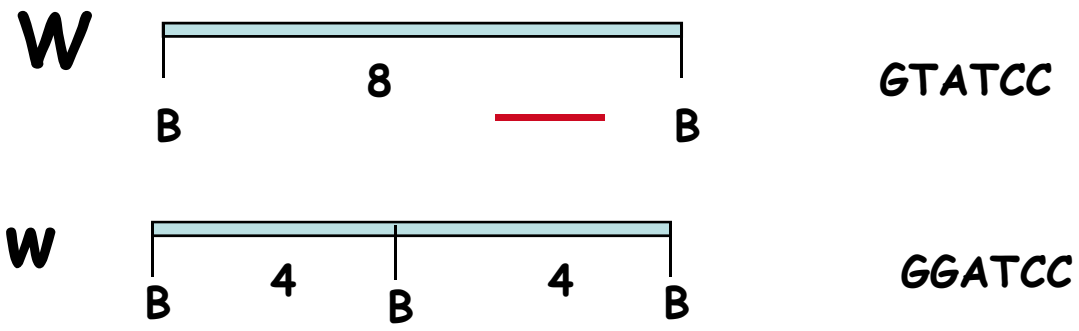
WB/wb	**x**	**wb/wb**
Female		**Male**
Wings		**No wings**
Bristles		**No Bristles**
Southern blot:		**Southern blot:**
5 and 2 kb band		**5 kb band**

There are some phenotypes for specific genes that are very painful to measure

Having a RFLP makes the problem easier

Mapping

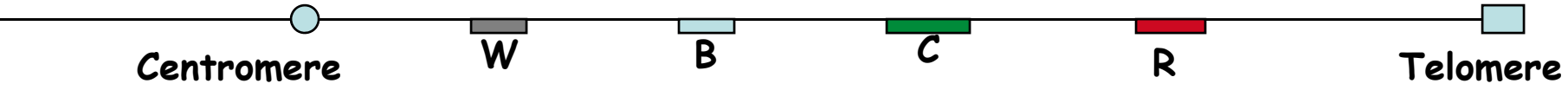
The same southern blot method can be employed for the (W) wing Locus with a different restriction enzyme (BamHI) if an RFLP exists at this locus !!
You make the DNA, digest half with EcoRI and probe with bristle probe
Digest the other half with BamHI and probe with the wing probe.



Mapping

To find the map distance between genes, multiple alleles are required.

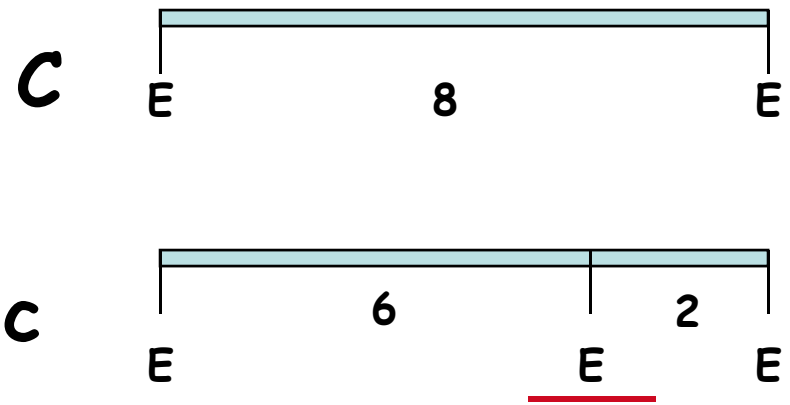
We can determine the distance between W and B by the classical Method because multiple alleles exist at each locus (W & w, B & b)



You find a new gene C. There are no variants of this gene that alter the phenotype of the fly, that you can observe. Say we don't even know the function of this gene. You can't even predict its phenotype.

However the researcher identified an RFLP variant in this gene.

Mapping



With this RFLP, the *C* gene can be mapped with respect to other genes:

Genotype/phenotype relationships for the W and *C* genes

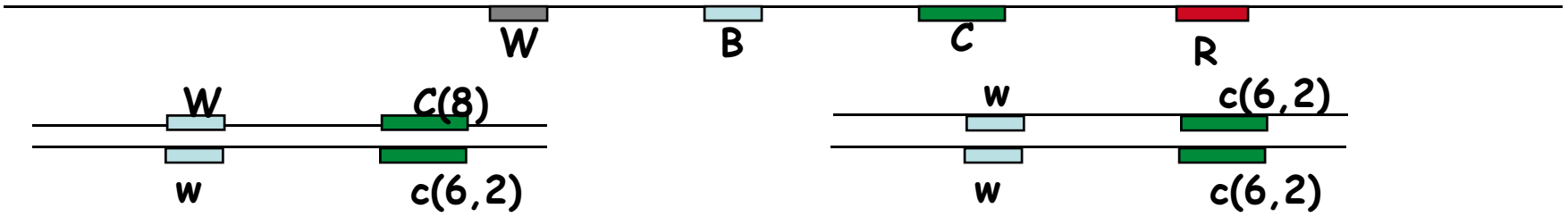
WW and Ww = Red eyes
ww = white eyes

***CC* = 8kb band**
***C/c* = 8, 6, 2 kb bands**
***cc* = 6, 2 kb bands**

To determine map distance between R and *C*, the following cross is performed

W C	**w c**
----------	**----------**
----------	**----------**
w c	**w c**

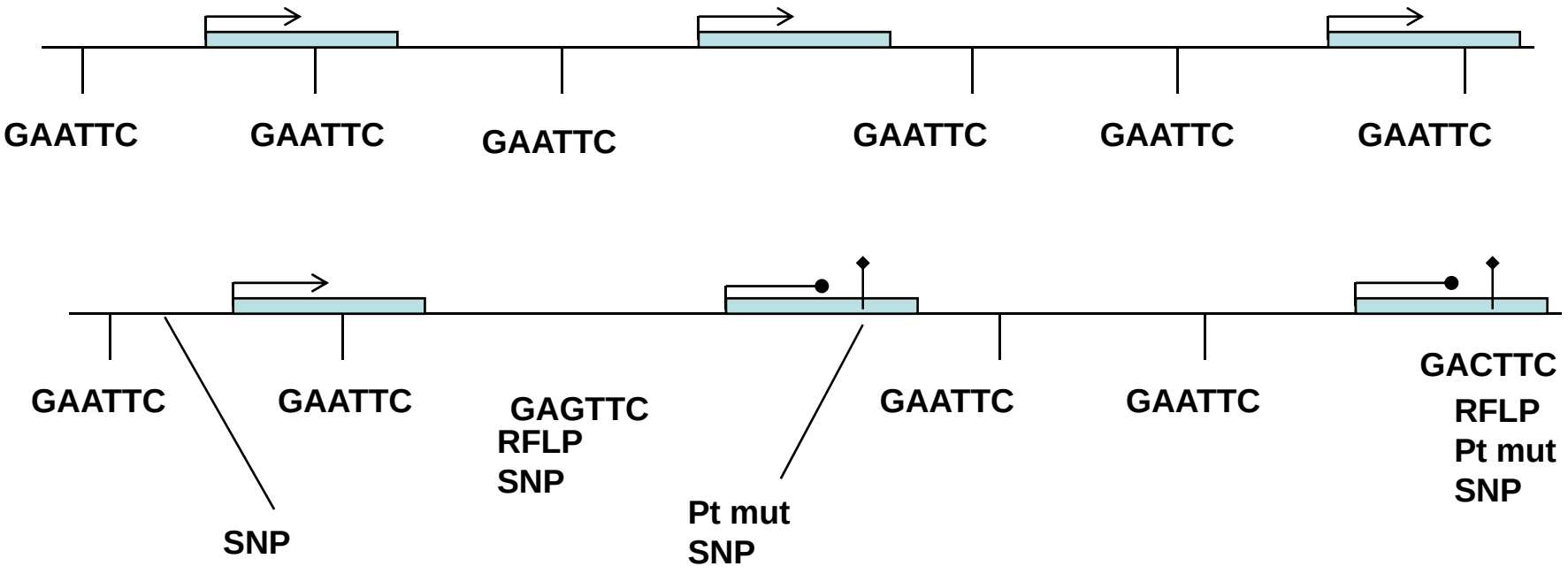
Mapping



Male gamete (wc)

Female gamete

SNPs, RFLPs, point mutations



PCR

If a region of DNA has already been cloned and sequenced, the sequence can be used to isolate and amplify that sequence from other individuals in a population.

Individuals with mutations in p53 are at risk for colon cancer

To determine if an individual had such a mutation, prior to PCR One would have to clone the gene from the individual of interest (construct a genomic library, screen the library, isolate the Clone and sequence the gene).

With PCR, the gene can be isolated directly from DNA isolated from that individual.

No lengthy cloning procedure

Only small amounts of genomic DNA required

30 rounds of amplification can give you $>10^9$ copies of a gene

Genotype and Haplotype

In the most basic sense, a haplotype is a “haploid genotype”.

Haplotype: particular pattern of sequential SNPs (or alleles) found on a single chromosome in a single individual

The DNA sequence of any two people is 99 percent identical.

Sets of nearby SNPs on the same chromosome are inherited in blocks.

Blocks may contain a large number of SNPs, but a few SNPs are enough to uniquely identify the haplotypes in a block.

The HapMap is a map of these blocks and the specific SNPs that identify the haplotypes are called tag SNPs.

Haplotyping: involves grouping individuals by haplotypes, or particular patterns of sequential SNPs, on a single chromosome.

Microarrays, and sequencing are used to accomplish haplotyping.

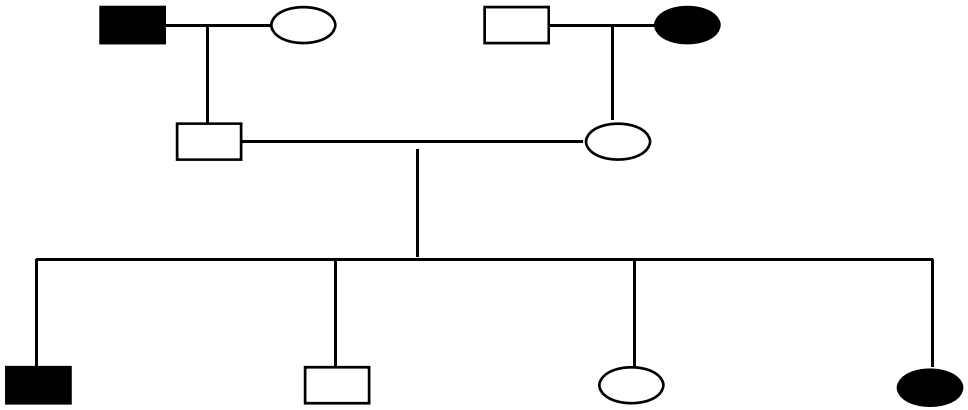
SNP mapping is used to narrow down the known physical location of mutations to a single gene.

The human genome sequence provided us with the list of many of the parts that make a human.

The HapMap provides us with indicators which we can focus on in looking for genes involved in common disease.

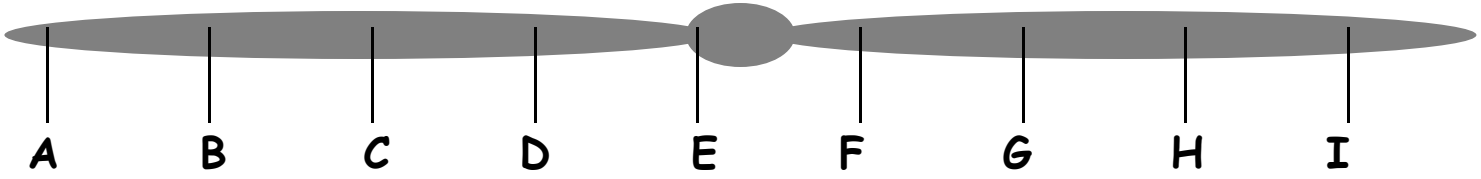
Using the HapMap data we compare the SNP patterns of people affected by a disease with those of unaffected people.
This allows researchers to survey the whole genome quickly and identify genetic contributions to common diseases--the HapMap Project has simplified the search for gene variants.

A recessive disease pedigree



Mapping recessive disease genes with DNA markers

DNA markers are mapped evenly across the genome
The markers are polymorphic- they look slightly different in
Different individuals.
We can tell looking at a particular individual which grandparent
Contributed a certain part of its DNA.
If we knew that grandparent carried the disease, we could say
That part of the DNA might be responsible for the disease.



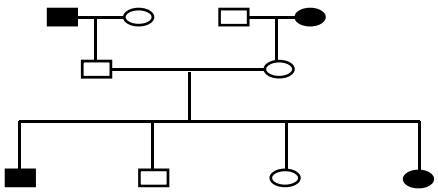
4 different alleles at each locus

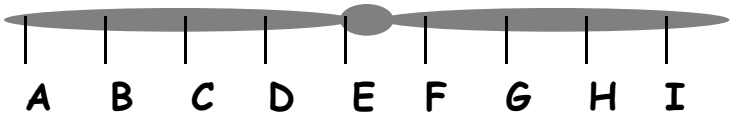
A1, A2, A3, A4

B1, B2, B3, B4

C1, C2,…….

Mapping recessive disease genes with DNA markers





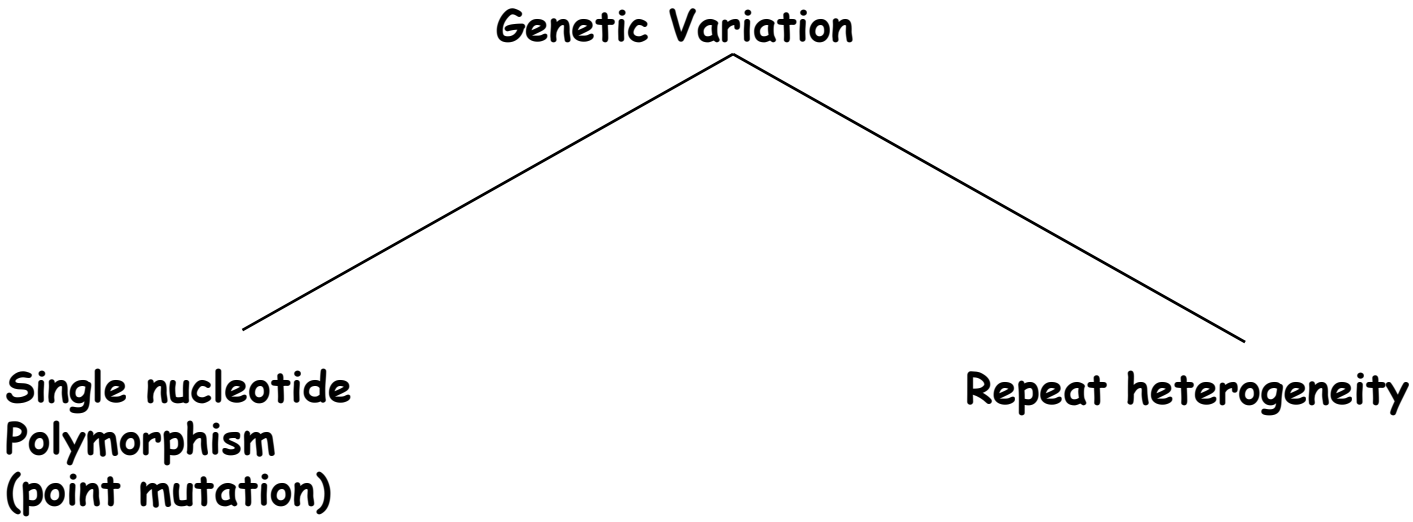
Grandparents 1 and 4 and offspring 1 and 4 have a disease
We would look at the markers and see that ONLY at position G do offspring 1 and 4 have the DNA from grandparents 1 and 4.
It is therefore likely that the disease gene will be somewhere near marker G.

Genetic polymorphism

•Genetic Polymorphism: A difference in DNA sequence among individuals, groups, or populations.

•Genetic Mutation: A change in the nucleotide sequence of a DNA molecule.

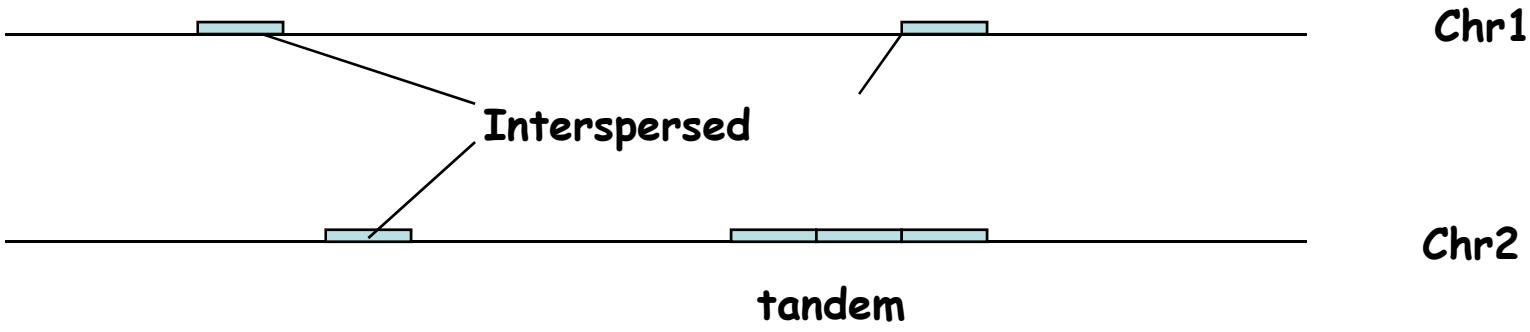
Genetic mutations are a kind of genetic polymorphism.



Repeats

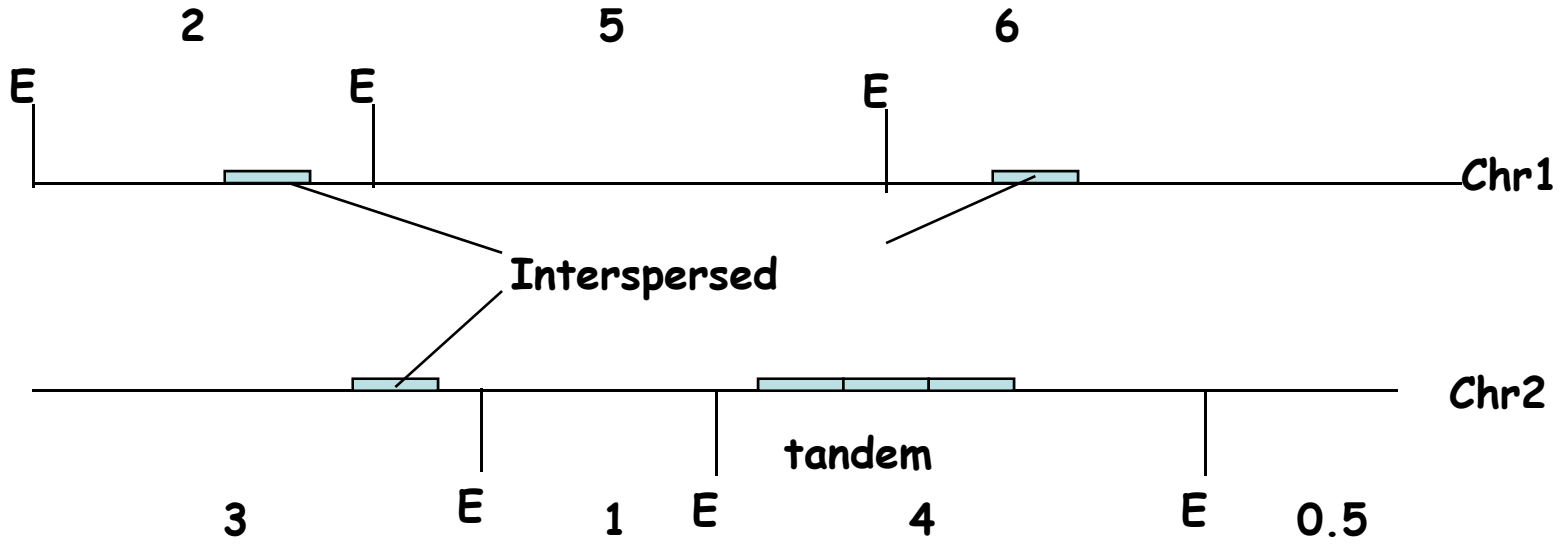
Variation between people- small DNA change – a single nucleotide polymorphism [SNP] – in a target site,
RFLPs and SNPs are proof of variation at the DNA level,

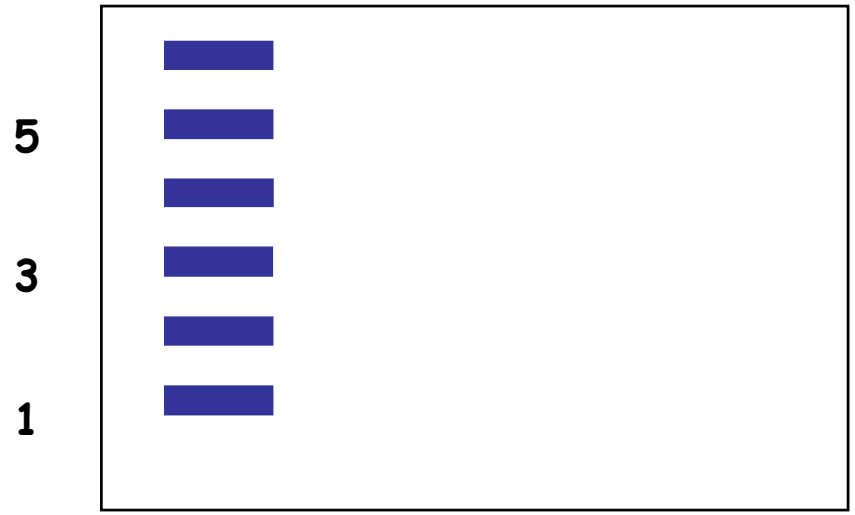
Satellite sequences: a short sequence of DNA repeated many times in a row.



Repeats

Satellite sequences: a short sequence of DNA repeated many times in a row.





Repeat probe

Repeat expansion

Tandem repeats expand and contract during recombination.
Mistakes in pairing leads to changes in tandem repeat numbers